Table 9. Summary Of SD-DI Features For Figure 22

Scene	TM-SD-DI	TSD-SD-DI	TRMS-SD-DI
DS1	18.1	5.9	19.1
1/2 DS1	18.4	5.7	19.3
1/4 DS1	22.0	4.6	22.5

3. CONCLUSIONS AND RECOMMENDATIONS

Objective feature extraction techniques have been presented that measure the predominant artifacts present in digitally transmitted video systems. Among these artifacts are blurring/smearing, blocking, edge busyness, image persistence, and jerkiness. Features are extracted from the digitized video imagery that reflect degradations perceived by the viewer. The features are sensitive to the type of video being transmitted which is important since the performance of digital codecs depend strongly on the type of video being transmitted. In addition, the features possess many of the desirable properties that humans also possess, including the potential adaptability to focus attention on local disturbances in the video. Thus, the features are expected to correlate strongly with subjective quality ratings.

Depending upon the feature one wishes to extract, the method for temporally aligning the input and distorted output video frames varies. Spatial blurring and jerkiness measures have been presented that do not require the input and output video scenes to be aligned. Other measures, such as edge busyness, blocking, image persistence, and jerkiness for natural motion scenes, require some form of temporal alignment. Two possible methods of temporally aligning the input and output video were presented. The computational requirements of the proposed features varied. However, these computational requirements appear reasonable for modern digital signal processing systems.

Spatial blurring features were presented that relate to the sharpness of the edges in the video imagery. These spatial blurring features appear to be applicable to many types of video imagery, including natural scenes. Blocking, edge busyness, and image persistence were shown to be forms of false edge energy appearing in the output video. Since the importance of edges are well recognized in the areas of human vision and object recognition, it is expected that these spatial blurring, blocking, edge busyness, and image persistence features will correlate especially well with subjective quality ratings.

Two measurement techniques for the jerkiness artifact of digitally transmitted video systems have been proposed. The temporal root mean square position error (TRMS-PE) feature compares the horizontal and vertical positions of a moving object in the output video scene to those in the input video scene. Although a ball was used for the moving object in the presented example, the technique is general enough to substitute any object. The TRMS-PE feature has been shown to be an accurate and repeatable measurement that determines the temporal positioning accuracy of a codec. Additional features were presented that measured the jerkiness of arbitrary video scenes. These included the missing frame ratio (MFR) feature and the TSD-SD-DI feature computed from the standard deviation of the error difference images. Video data compression is often a tradeoff between allocating bits between temporal positioning accuracy and spatial resolution. The ability to measure separately these two attributes raises the possibility of tailoring performance specifications to the application.

Further work needs to be done to determine the optimal method for combining all of the extracted feature values to produce an overall quality rating (the quality classification subsystem shown in Figure 1). Properly combining the many feature measurements into an overall quality assessment rating may require an understanding of the temporal and spatial properties of the eye and brain. To be universally useful, the quality classification subsystem must perform well over a wide range of applications. To obtain this goal, the quality classification system may require user specific application information. Subjective test results on imagery that spans the full range of digitally transmitted video systems should be used to select an optimal set of features, to train the quality classification subsystem, and to evaluate the performance of the completed system.